DOCUMENT RESUME

ED 479 825 JC 030 402

AUTHOR	Luan, Jing
TITLE	Developing Learner Concentric Learning Outcome Typologies Using Clustering and Decision Trees of Data Mining.
PUB DATE	2003-05-00
NOTE	16p.; Paper presented at the Annual Forum of the Association for Institutional Research (43rd, Tampa, FL, May 18-21, 2003).
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	EDRS Price MF01/PC01 Plus Postage.
DESCRIPTORS	Classification; *Cognitive Processes; Community Colleges; Learning; *Learning Processes; *Outcomes of Education; Student Centered Curriculum; *Student Characteristics; Two Year College Students; Two Year Colleges

ABSTRACT

This study aims to address learning outcomes from the perspective of learners. The research questions asked were: (1) What learner concentric ideas can be used to indicate the outcomes of learning? (2) How would the new learning outcome index be used to generate typologies? (3) What are the inner relationships of the typologies? and (4) How can the typologies be applied? The author argues that typology is fundamental to science, yet it is seriously underused and under-researched in social science. Only a handful of authors have worked on the subject, which has created an insurmountable gap between what has been done and what needs to happen. The study examined first-time college students enrolled in spring 1996 at a suburban college on the West coast with enrollment of 15,000 per semester. The study tracked the students for 6 years for their enrollment behavior, graduation status, and transfer status. The study theorizes that what may be important to the learner may not be important to the institution. The author argues that the "big three" (gender, age, race) are not good predictors of typologies, and that learners' behaviors as reflected by the OIndex are a better way of describing learner outcomes. (Contains 11 references.) (NB)

Reproductions supplied by EDRS are the best that can be made
from the original document

ED 479 825

Developing Learner Concentric Learning Outcome Typologies Using Clustering and Decision Trees of Data Mining

Jing Luan

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

- This document has been reproduced as received from the person or organization originating it.
- Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

J. Luan

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

JC 030 402

BEST COPY AVAILABLE

Developing Learner Concentric Learning Outcome Typologies Using Clustering and Decision Trees of Data Mining

Jing Luan, Ph.D.
Chief Planning, Research & Knowledge Systems Officer
Cabrillo College

Rationale

There is an explosion of scholarly work in the area of *learning* outcomes in recent past as judged by the amount of search hits in ERIC, the number of authors and the enumeration of new terms, such as Incidental Learning (Mealman, 1993) and Asynchronous Learning (Browne, et al, 2000). Just in the last five years, learning outcomes articles increased by close to 40%. The field of research quickly shrinks when the dimension of student-centered outcomes is taken into the equation. Spadano (1997) provided philosophical analysis of contrasting pedagogical approaches of either teacher- or student-centered. Lee (1999) discussed student-centered problem based learning (PBL). Scarcity of research or managerial attention has been paid to *learner* outcomes and the assessment of which. This study attempts to address the "learning outcomes" – a term that has stuck – from the perspective of learners. Specifically, using available data in a college data warehouse, it developed learner behavior based outcome measures, collectively called the OIndex, and further classified learners based on the OIndex. The results of the classificatory work manifested as learning outcome typologies.

Typology is fundamental to science (Bailey, 1994; Fenske, et al., 1999) and is seriously underused and under-researched in social science (Luan, 2002). Astin (1993) conducted an empirical typology of college students in hopes of gaining some insights into student life. Fenske et al. (1999) proposed an early intervention program typology. Levine et al. (2001) developed an empirically based typology of attitudes toward learning community courses. All in all, only a handful of authors have worked on this subject, which has created an insurmountable gap between what has been done and what needs to happen.

Technically, the commonly accepted view of a set of typologies for a particular subject area refers to its members or entities in a group that are maximally similar and members between groups maximally dissimilar. The more distinct the groupings are the better the typologies. Differences are mathematically driven. They are either defined by the distance measure $D = \sum (x_{Ai} - x_{Bi})^2$, such as clustering algorithms or by correlational measures (R-analysis), such as factor analysis.

Typologies are not student profiles seen in fact books or enrollment reports. In addition, the common use of students' characteristics that are outside the control of the institution or the student such as gender, age, or race – the *big three*, continues to put the institution and the student at odds with each other. The unintended consequence of using the *big*

TC030402

three is the unfortunate acceptance, or perceived acceptance, of causal-effects. Students belonging to particular race or ethnic groups are frozen in time, stuck in the perceived behavior patterns. It is least scientific or socially responsible to allow "if you belong to this race, you then will do this." As common sense would dictate, persons sharing the same *big three* most often behave differently. One benefits of the study, perhaps, is to shift the focus of attention to behaviors of learning and away from biological features of students.

Typologies already exist in higher education institutions. A university or community college mission statement typically describes the type of courses they offer and the type of students they serve. These are qualitative typologies that help describe ***what*** they do. On the other hand, true typologies are for the purpose of describing ***how*** they will do it. In this manner, typologies classify events, activities and persons associated with a particular mission into distinct groupings for the purpose of providing customized intervention. With typological groupings, everyone is equal, but also special.

The purpose of monitoring learning outcomes is to improve an institution's effectiveness. For this reason, the deeper we understand our learners the better the improvements. The following graph illustrates the typical practice in the private sector to maximize their market shares through behavior based analysis of their customers.

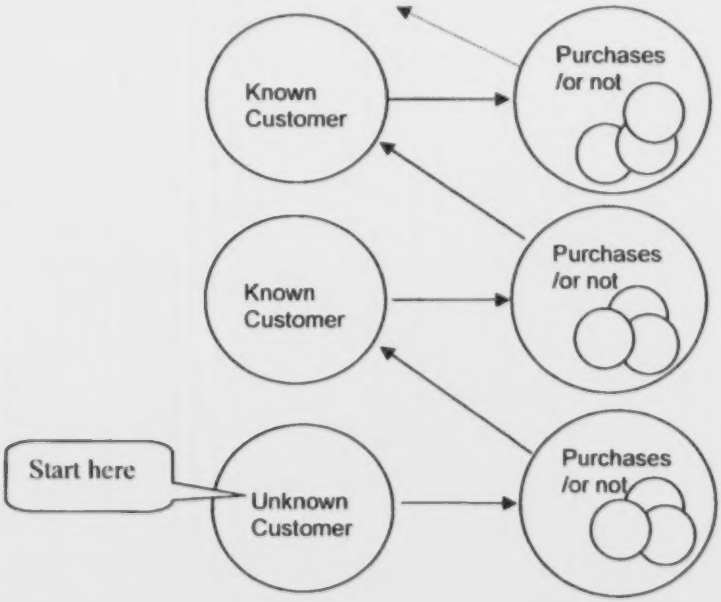


Figure 1. Spiraling phases of customer centered model

This is a customer centered model. It starts from the moment the customer initiates a connection, purchases goods all the way to the customer returning for more. When the first transaction is complete, the customer's status changes from unknown to known. The company researchers study the data collected from the purchases to classify customers

into low users, browsers, returned customers, etc and to develop strategies to either meet the needs of the customers or increase their satisfaction or a combination of the two. It continues as a spiraling loop. This model, although highly simplified, helps emphasize the point that customer behaviors are paramount and companies work to satisfy the customers. Students are not customers and higher education is not in the business of selling products. However, the principles of understanding customers are applicable in higher education. The task at hand is to find ways to truly make our learning outcomes research to be learner concentric (learner centered) and to identify effective methods to classify our learners.

Literature search produced little on student or learner centered learning outcomes assessment. The existing learning outcomes are designed *by* the institution and *for* the institution. Institutions approach learning outcomes by what the provider believes ought to be learned. Traditional measures have been generally regarded as inappropriate, such as degrees, grades. Employment data, although most useful as independent evidence, are difficult to come by. What then is available to be used to measure learning outcomes?

There is business education, but not education business. A great number of modern education practitioners do not consider learning to be a business transaction. The author believes that learning is a *process* with a great amount of time, financial investment and most importantly psychological commitment (Luan, 2001). Learning is one of the most complicated transactions of all. Therefore, it behooves researchers to examine the learning from the angle of a learner. What the learner sees may not be what the provider of learning sees. The way the learner perceives learning may provide additional information to improve current education decision making.

The graph below illustrates the availability of student data for understanding learners at a typical community college, which happened to be the site chosen for this research. The service usage data is either incomplete or entirely absent from the database. However, the institution has a well defined MIS database of course related data. The box on the right contains the typical outcome measures designed and monitored by the institution. The institution is moving toward course based learning outcomes in which learners and professors clearly state their desired course objectives. At least one learning outcome institute has been in place for three summers. The data from classroom based learning outcomes are outside the scope of this research.

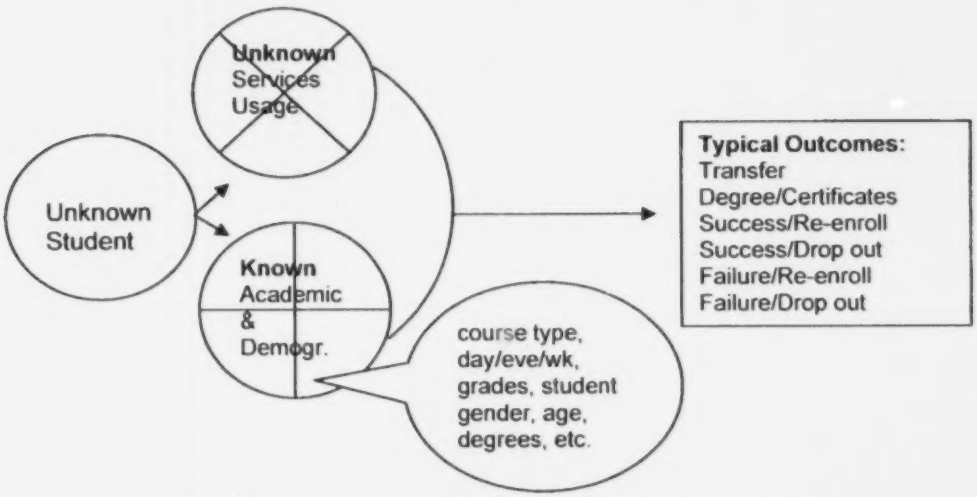


Figure 2. Empirical data modeling for typological research of the study

This research chose to focus exclusively on the academic data in the data warehouse to develop a learner concentric outcome index and pull in learner demographics to assist with understanding of the typologies.

Research Questions:

1) What learner concentric fields can be used to indicate the outcome of learning?
2) How would the new learning outcome index be used to generate typologies?
3) What are the inner relationships of the typologies?
4) How can the typologies be applied?

Design:

The underpinning frame of thought for developing learner concentric outcome indexes is as follows. Financial, social, family obligations, and psychological readiness, are among a number of influence factors a student uses to determine his/her postsecondary education (Hossler, 1984). The question is what are the results of the influence factors? The answer lies in the congruent interplay of how many courses a learner thinks s/he can reasonably take and what s/he would like to study. Therefore, for a learner, the number of courses and the potential academic requirements are weighed equally and simultaneously.

Outcomes are semester based. To a learner, learning is a purposeful process broken down into semesters. S/he approaches learning on a semester by semester basis. Many factors that do not necessarily attract the attention of the institution are important to the learner. For example, their family obligations, employment status, financial capacity, distance to college, stage in life, job requirements, and offerings of the colleges all come into play. Most of these factors manifest themselves as the types of courses they take, the number

BEST COPY AVAILABLE

of courses they take, the time they take them. These are the choices they make. This is the first dimension. It is manifested in the number of courses they take, called Course Volume in this study, and the types of course they take. The second dimension is their expectations of success in the class, which is assumed to be the grade of an "A". They most likely want to succeed in each of the classes they choose. This dimension is considered to be the "ideal outcome". At the end of the semester, outcomes will emerge that more often than not are different from their desired outcome. This is the gap, or distance to goal. Learners all make the best efforts to achieve their desired outcome. They will put forward enough efforts till they reach their threshold at which their option (strategy) is to withdraw from class. Even though the detailed reason why a particular learner withdraws from a class remains a long lasting research and academic debate, the fact that a learner withdraws from a class means s/he is reacting to something. When someone reacts by withdrawing, the person is making adjustments to his/her studies. This action of withdrawing is considered the Adjustment Factor.

The study defines learning outcomes as the outcomes of learner concentric course taking behaviors with the semester as the unit of time. Specifically, the Outcome Index (OIndex) is the congruent use of three separate indexes. They are distance to goal, the adjustment factor and course volume.

Distance to Goal (D2G) = Real Grade Points (RealGPt) – Ideal Grade Points (IdealGPt),
where,

$$\text{RealGPt} = \sum \text{(Units x GPt) + (Units x GPt)...}$$

$$\text{IdealGPt} = \sum \text{(Units x GPt) + (Units x GPt)* ...}$$

Course Volume (CrsCnt) = Count of courses taken

$$\text{Adjustment Factor (AFactor)} = \left\{ \sum \frac{Ws}{CrsCnt} \right\} \text{x } 100$$

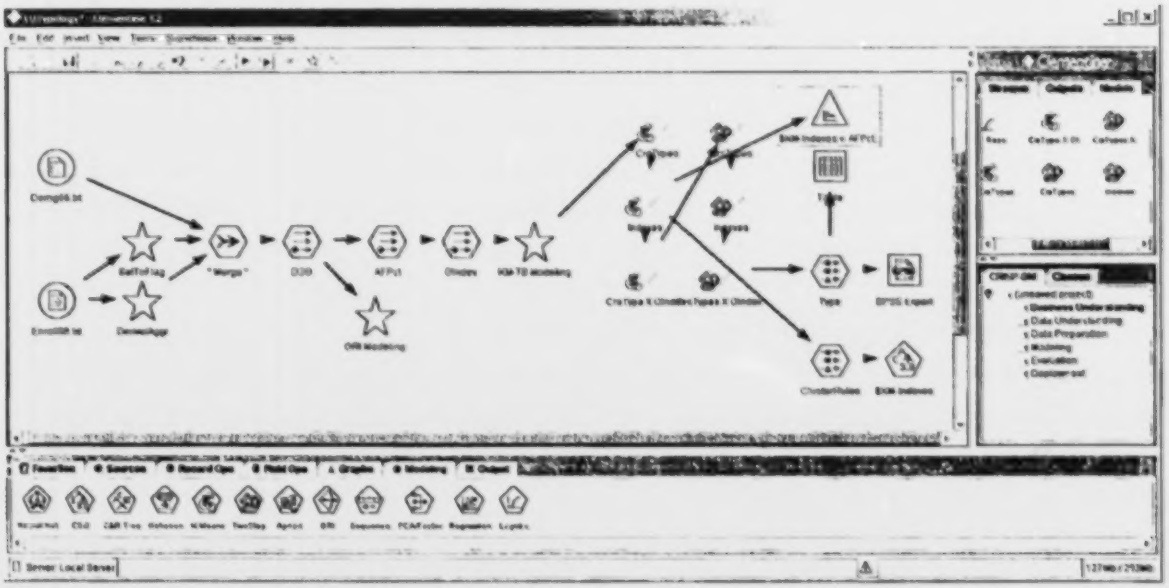
* Note: Grade points (GPt) for Pass/Fail courses have been adjusted.

The OIndex, a group of three sub-indexes, is considered to be the primary candidate for clustering. The study identified two additional candidates for clustering. One of them is the varieties of course types. They could be transfer courses, basic skills courses, vocational education courses, courses for leisure (please contact the author for detailed taxonomy of these courses). The third candidate was simply the combination of the first two – a meta group.

The study chose to examine first-time college students enrolled in spring 1996 at a suburban community college on the west coast with 15,000 enrollment per semester. The study tracked these students for six years for their enrollment behavior, graduation status and transfer status. The study used unsupervised clustering algorithms of K-Means and TwoStep to generate the groupings, which were evaluated for their centroid distance (inter-group and intra-group). To understand the inner workings of the clusters produced by K-Means and TwoStep, the study used C5.0, a decision tree based algorithm, to examine the split of the branches. To the extent possible, the study examined the traditional outcome measures of graduation, transfer, etc.

Method - Database Development and Data Mining

The study first used Brio Query for cohort identification due to the fact that the author had existing java codes written for similar projects in Brio. The results from Brio were exported as .txt files that were directly accessible to Clementine for the calculation of new fields, conducting clustering and decision tree analyses. When necessary, Clementine exported data into SPSS 11.0 for 3-D graphical analysis of clusters. The following screen shot illustrates the data stream built within Clementine for the entire clustering study, including the nodes used for calculating fields and the decision tree algorithm. The power of Clementine is enormous. Among them, the ability to directly interface with static or live databases, to calculate new fields using GUI guided nodes, to convert transactional data file into analytical data file and to allow infinite number of scenarios being built and examining using the 12 different algorithms. Everything is done on one stream, which makes it so much easier for cross-validation and documentation.



Generating Typologies

Generating typologies requires a good understanding of the input variables (fields). Typically 7 plus or minus 2 dimensions are considered appropriate. However, if there are more fields, it is advisable to consider auxiliary analyses to eliminate fields highly correlated with other fields or group certain fields into "dimension groups".

Clementine 7.2, a powerful workbench with industrial level analytical capacity, was used for the study for the data manipulation tasks of converting rows to columns (transactional data to analytical data transformation, also called case to var) and calculation (deriving new fields) of new fields.

BEST COPY AVAILABLE

Findings

The study used clustering algorithms of K-Means and TwoStep. The study took advantage of the scenario building feature of Clementine, which allows infinite number of scenarios for each of the modeling algorithms to be built on the same workbench, so that the researcher can compare and contrast the scenarios and pick the best one. Since the author chose three groups of fields as clustering candidates, the following matrix describes the number of scenarios built by Clementine:

Table 1: Clustering Scenario Matrix:

	Course Type	OIndex	Course Type + OIndex
K-Means	3, 4, 5	3, 4, 5	3, 4, 5
TwoStep	3, 4, 5	3, 4, 5	3, 4, 5

With each algorithm tested for 3, 4 and 5 clusters separately for Course Type, OIndex, and Course Type + OIndex, the study built and examined a total of 18 scenarios. Every one of the scenarios was examined using 3-D data visualization technique as well as Chi-square tabulation of demographics and other related fields.

Of all 18 scenarios, K-Means produced one optimal cluster scenario with the best inter-cluster separation as illustrated by the 3-D graph below.

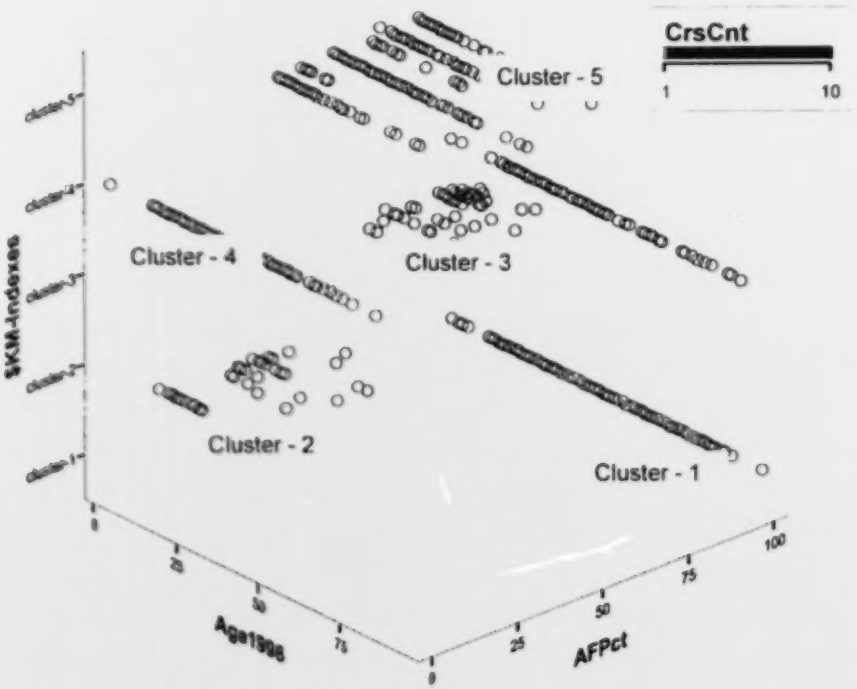


Figure 3. 3-D rendition of the clusters by age and by Adjustment Factor

BEST COPY AVAILABLE

The optimal cluster resulted from the use of the OIndex of fields. The TwoStep algorithm also confirmed this with its own slightly different arrangement of cluster memberships. However, K-Means' scenario is by far the best. A decision tree below helps describe the logic of the cluster separation.

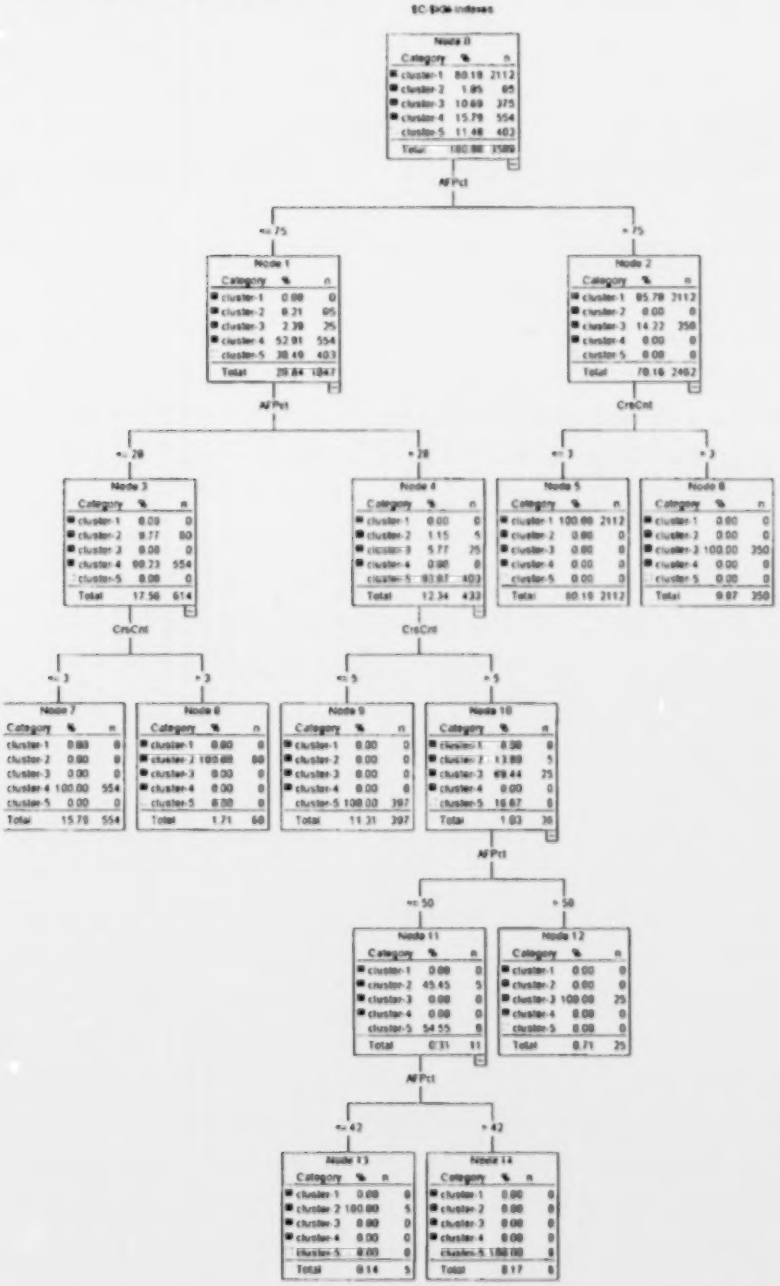


Figure 4. Binary split of decision tree to describe the cluster separation

BEST COPY AVAILABLE

Cluster membership:

Cluster-1: 2,112
Cluster-2: 65
Cluster-3: 375
Cluster-4: 554
Cluster-5: 403

One rule for determining the adequacy of cluster membership states that the smallest cluster membership should be more than 20% of the size of the largest cluster membership. The formula is (Smallest Cluster/Largest Cluster)*100. Apparently this rule is violated using the formula: (65/2112)*100 = 3.07%. However, the purpose of this study is to research into the behaviors of learners, therefore, smaller clusters may denote outliers that are as important as others that fall within the Gaussian distribution.

The study conducted a Mean Analysis using the three sub-indexes of the OIndex and clusters. Course Volume has been multiplied by 10 so that it is not drowned out by the other two indexes.

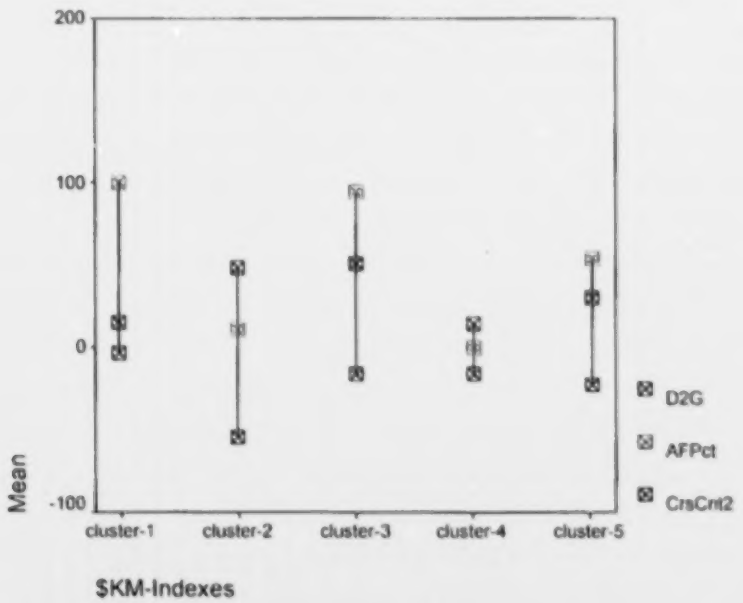


Figure 5. Drop-line analysis of the clusters by OIndex

The drop-line chart helps describe the differences of the clusters based on the mean of the three sub-indexes. Cluster-1 has high Adjustment Factors, lowest D2G and reasonable Course Volume. Cluster-2 has the highest D2G, but very low Adjustment Factor. Cluster-3 is similar to Cluster-2 except it has high Course Volume. Cluster-4 is not extraordinary in any manner. Cluster-5 has the second highest D2G, but middle of the road Adjustment Factor and Course Volume. The following table denotes the preliminary naming of these clusters.

Table 2: Naming Clusters in the Sequence of the Clusters

	Cluster Mean Analysis Results			Name
	D2G	AFactor	Course Vol	
Cluster-1	Low	High	Low	Smart Adjustors
Cluster-2	High	Low	High	Disenchanted Heavy Loaders
Cluster-3	Medium	High	High	Heavy Loaders
Cluster-4	Medium	Medium	Medium	Samplers
Cluster-5	High	Medium	High	Back Breakers

The study theorizes that what is important to the learner may not be important to the institution. The following table attempts to rank the clusters in the order of the most desired types both by the learners themselves and a higher education institution.

Table 3: Hypothetical Ranking of Learner Typologies by Learner and by Institution

For The Learner (Most Desired First)	For The Institution (Most Desired First)
1. Smart Adjustors	1. Heavy Loaders
2. Samplers	2. Back Breakers
3. Heavy Loaders	3. Smart Adjustors
4. Back Breakers	4. Disenchanted Heavy Loaders
5. Disenchanted Heavy Loaders	5. Samplers

The Five Typologies by Demographics and Graduation:

The following three tables display the observed and expected cell counts of clusters and their demographics and graduation (defined by receiving at least one award).

Table 4. KM-Indexes and Gender Crosstabulation

		Gender			
		F	M	Unknown	Total
cluster-1	Count	1239	873	0	2112
	Expected Count	1185.7	925.7	0.6	2112
	% within $KM-Indexes	58.7%	41.3%	0.0%	100.0%
cluster-2	Count	35	30	0	65
	Expected Count	36.5	28.5	0	65
	% within $KM-Indexes	53.8%	46.2%	0.0%	100.0%
cluster-3	Count	201	174	0	375
	Expected Count	210.5	164.4	0.1	375
	% within $KM-Indexes	53.6%	46.4%	0.0%	100.0%
cluster-4	Count	279	274	1	554
	Expected Count	311	242.8	0.2	554
	% within $KM-Indexes	50.4%	49.5%	0.2%	100.0%
cluster-5	Count	216	187	0	403
	Expected Count	226.2	176.6	0.1	403
	% within $KM-Indexes	53.6%	46.4%	0.0%	100.0%
Total	Count	1970	1538	1	3509
	Expected Count	1970	1538	1	3509
	% within $KM-Indexes	56.1%	43.8%	0.0%	100.0%

There appears to have no major gender differences across all five clusters. This further reinforced the earlier statement that demographic variables, such as gender, may have no discernable differences across the typologies, even though the behaviors of the members across the clusters are most diverse.

Table 5 shows a high count of African American students in Cluster-2 - the Disenchanted Heavy Loaders characterized by high D2G, low Adjustment Factor, and high Course Volume. Is being African American responsible for being a disenchanted heavy loader? The answer lies in the fact that the OIndex, which generated the clusters (typologies) used purely learner behavior data. The small number of the cell counts led the author to believe that perhaps certain student culture or study groups may be at play. Regardless, the finding signals an important task to perform. The task is to advise these learners that they should either reduce their overall load or utilize their adjustment factor more.

Table 5: KM-Indexes and Race Crosstabulation

		Race								
		Af. Am.	Am. Ind.	Asian	Filipino	Hispanic	Other	Unknown	White	Total
cluster-1	Count	22	25	70	15	325	21	8	1626	2112
	Expected Count	29.5	27.7	72.2	26.5	346.7	25.3	9	1575.1	2112
	% within $KM-Indexes	1.0%	1.2%	3.3%	0.7%	15.4%	1.0%	0.4%	77.0%	100.0%
cluster-2	Count	6	0	4	1	6	2	1	45	65
	Expected Count	0.9	0.9	2.2	0.8	10.7	0.8	0.3	48.5	65
	% within $KM-Indexes	9.2%	0.0%	6.2%	1.5%	9.2%	3.1%	1.5%	69.2%	100.0%
cluster-3	Count	6	6	11	9	62	3	1	277	375
	Expected Count	5.2	4.9	12.8	4.7	61.6	4.5	1.6	279.7	375
	% within $KM-Indexes	1.6%	1.6%	2.9%	2.4%	16.5%	0.8%	0.3%	73.9%	100.0%
cluster-4	Count	8	5	19	10	107	6	3	396	554
	Expected Count	7.7	7.3	18.9	6.9	90.9	6.6	2.4	413.2	554
	% within $KM-Indexes	1.4%	0.9%	3.4%	1.8%	19.3%	1.1%	0.5%	71.5%	100.0%
cluster-5	Count	7	10	16	9	76	10	2	273	403
	Expected Count	5.6	5.3	13.8	5.1	66.2	4.8	1.7	300.6	403
	% within $KM-Indexes	1.7%	2.5%	4.0%	2.2%	18.9%	2.5%	0.5%	67.7%	100.0%
Total	Count	49	46	120	44	576	42	15	2617	3509
	Expected Count	49	46	120	44	576	42	15	2617	3509
	% within $KM-Indexes	1.4%	1.3%	3.4%	1.3%	16.4%	1.2%	0.4%	74.6%	100.0%

Table 6 found that Cluster-3 had significant more learners who graduated (received at least one award within 6 years). Recall Cluster-2 was the Heavy Loaders group characterized by medium D2G, high Adjustment Factor and high Course Volume. As a matter of fact, their transfer rate was also high (data not shown). Why did heavy loaders have high graduation rate? This may be due to the fact that institutions tend to favor those who take on heavy load and may have subconsciously pampered them (see Table 3, hypothetical ranking). It may also be a result of lack of service usage data for the study.

Table 6: KM-Indexes and Awards Crosstabulation

		Awards		
		N	Y	Total
cluster-1	Count	2027	85	2112
	Expected Count	2022.9	89.1	2112
	% within $KM-Indexes	96.0%	4.0%	100.0%
cluster-2	Count	65	0	65
	Expected Count	62.3	2.7	65
	% within $KM-Indexes	100.0%	0.0%	100.0%
cluster-3	Count	332	43	375
	Expected Count	359.2	15.8	375
	% within $KM-Indexes	88.5%	11.5%	100.0%
cluster-4	Count	548	6	554
	Expected Count	530.6	23.4	554
	% within $KM-Indexes	98.9%	1.1%	100.0%
cluster-5	Count	389	14	403
	Expected Count	386	17	403
	% within $KM-Indexes	96.5%	3.5%	100.0%
Total	Count	3361	148	3509
	Expected Count	3361	148	3509
	% within $KM-Indexes	95.8%	4.2%	100.0%

Discussion - Potential Implications of the Five Typologies

The Big Three is not a good "predictor" of typologies. Learners' behaviors as reflected by the OIndex are a better way of describing the learning outcomes. In addition, several major benefits can be obtained from using the OIndex by education institutions. Monitoring the sub-indexes, the Adjustment Factor in particular, will give the institution an early warning of trouble spots. Tracking of the OIndex will provide the institution a chance to set benchmark and to drill down to problem areas if needed. For example, when a particular sub-index shifts, decision makers would like to know which cluster attributed to the change and which subgroup of the cluster is accounted for the change. By not casting a wide net, the institution can run a more focused and effective campaign of any kind as guided by the OIndex. By associating sub-indexes to student success, the institutions can better counsel learners on their course taking patterns. By adding the predicative modeling, the institutions can forecast retention rates for the entire institution and for an individual learner. The institution can use the movements of the OIndex and the affiliated clusters to link to bigger climate of external factors of the economy and population change, to name a few. Lastly, it is a field an institutional research professional can muster a quick win by using their data warehouses and their specialized statistical skills.

Future Research

Conduct predicative modeling for retention or GPA using homogenized clusters. Inclusion of service usage data would theoretically enhance the predictability power of the clusters. Inclusion of student satisfaction data would further increase the comprehensiveness of the learners. Also helpful is to include electronic portfolio data.

Work should be continued to examine how to apply OIndex to measuring learning outcomes.

Conclusion

The use of the typologies in this study will help illustrate the importance of learner concentric philosophy of educational practice; better understand how learners' behaviors can influence institutional outcomes, and ultimately institutional effectiveness; provides potential areas for interventions by zeroing on the specific needs of learners who belong to different typologies; increase organizational efficiency; providing competitive advantage; better respond to the changing environment and education market demands.

Data are produced not by the institution, but by the learners. Let data from the learners speak for themselves and not being dictated by the institution.

Contacting the author

Jing Luan, Ph.D.
Chief Planning, Research & Knowledge Systems Officer
Cabrillo College
6500 Soquel Drive
Aptos, CA 95003
831.477.5656
[jing]@cabrillo.edu
(Do not use brackets when sending him email. Brackets prevent automatic email harvesting)

Bibliography

Astin, A. (1993). An empirical typology of college students. Journal of College Student Development, 34, 36-46.

Bailey, K. (1994). Typologies and taxonomies: An introduction to classification techniques. (Sage University Paper series on Quantitative Applications in the Social Sciences, series no. 07-102). Thousand Oaks, CA: Sage.

Browne, J., Warnock, S., & Boykin, N. (1997) Work-related Outcomes For Instructors Using Asynchronous Learning Networks. ERIC_NO: ED414187

Fenske, R., Keller, J., & Irwin, G. (1999). Toward a Typology of Early Intervention Programs. Advances in Education Research. Vol. 4.

Hossler, D. (1984) Enrollment Management – an integrated approach. The College Board. New York, New York.

Lee. J. (1999) Problem-Based Learning: A Decision Model for Problem Selection. Proceedings of Selected Research and Development Papers Presented at the 21st National Convention of the Association for Educational Communications and Technology, Houston, Texas. ED436162

Levine, J. Jones, P. & Williams, R. (2001) Developing an Empirically Based Typology of Attitudes Toward Learning Community Courses. Part of Presentation for the AAHE Assessment Conference. Denver, Colorado

Luan, J. "Learner Relationship Management: Knowledge Management and Data Mining as Applied in Higher Education." Paper presented at the Annual Conference of the Community College League of California, Riverside, CA, 2001.

Luan, J. (2002) Mastering Data Mining: Predicative Modeling and Clustering Essentials. AIR Forum Workshop Manual. AIR 2002. Toronto, Canada

Mealman, C. (1993) Incidental Learning by Adults in a Nontraditional Degree Program: a Case Study Presented at the Midwest Research-to-Practice Conference in Adult, Continuing, and Community Education, Columbus, Ohio, October 13-15, 1993

Spadana, J. Zeidler, D. L., & Chappell, M., (1997) Advancing Ownership of Understanding and Responsibility through Homework in Mathematics Education. Paper presented at the 40th Annual Meeting of the Pi Lambda Theta International Education Leadership Association, San Diego, California. ED414187.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title:	Developing Learner Concentric Learning Outcome Typologies Using Clustering and Decision Trees of Data Mining
Author(s):	Jing Luan, Ph.D.
Corporate Source:	Presentation at AIR (Association for Institutional Research) Institutional Research) Conference Tampa, FL., 2003
Publication Date:	May, 2003

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 1	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2A	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2B
Level 1 ■	**Level 2A** ☐	**Level 2B** ☐
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) and paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only.	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only.

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Signature:		Printed Name/Position/Title:	Jing Luan, Ph.D., Chief Planning, Research & Knowledge Sys tems Ofc
Organization/Address:	Cabrillo College 6500 Soquel Dr., Aptos, CA 95003	Telephone:	831-477-5656
		FAX:	831-477-5250
		E-Mail Address:	jing@cabrillo.edu
		Date:	8/13/03

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
4483-A Forbes Boulevard
Lanham, Maryland 20706

Telephone: 301-552-4200
Toll Free: 800-799-3742
FAX: 301-552-4700
e-mail: info@ericfac.piccard.csc.com
WWW: http://ericfacility.org